J. Child Lang. **41** (S1) (2014), 48–63. © Cambridge University Press 2014
doi:10.1017/S0305000914000282

First language development: a usage-based perspective on past and current research*

ELENA LIEVEN

School of Psychological Sciences, University of Manchester

ABSTRACT

I first outline three major developments in child language research over the past forty years: the use of computational modelling to reveal the structure of information in the input; the focus on quantifying productivity and abstraction; and developments in the explanation of systematic errors. Next, I turn to what I consider to be major outstanding issues: how the network of constructions builds up and the relationship between social and cognitive development and language learning. Finally, I briefly consider a number of other areas of importance to a psychologically realistic understanding of children's language development.

The forty years in which *JCL* has been publishing have seen major developments in research on child language. These include new theoretical approaches, important advances in methodology, and a plethora of empirical studies. In what follows, I highlight what I see as major developments in understanding the first language development of typically developing children, with some examples. I then raise a number of issues which need considerable further research. Inevitably, I approach this from the usage-based theoretical framework within which I work. The fundamental proposals of this approach are that children learn language from particular utterances in particular contexts and that increasingly complex and abstract linguistic representations are built from this learning. The major issues are, therefore, what information there is in the input and how children extract it; the role of domain-general processing in language learning; and explaining why children's language does not always look like that of adults.

MAJOR DEVELOPMENTS

Computational modelling: discovering the information in the input

Children learn language from what they hear as this interacts with their domain-general cognitive learning processes, the current state of their

[*] Address for correspondence: Elena Lieven, School of Psychological Sciences, University of Manchester, M13 9PL, UK. e-mail: elena.lieven@manchester.ac.uk

https://doi.org/10.1017/S0305000914000282 Published online by Cambridge University Press

language system, and their communicative intentions. For the usage-based approach, the structure of the input and what children can extract from it is crucial. The collection and transcription of large corpora of children's and their caretakers' speech combined with the development of more sophisticated computational techniques has been very important in demonstrating how much information is available in the input and how this might impact on learning.

One of the first studies on the extraction of lexical categories was by Redington, Chater, and Finch (1998). In this study, words were clustered together on the basis of the number of contexts they shared (defined as the preceding and following word). This yielded clusterings of words that reflected lexical categories with reasonable accuracy. Using a rather different approach of extracting frequent lexical frames of two words with a slot between them, Mintz (2003) showed that for English the frames tended to gather words of the same lexical category in the slots (either nouns or verbs), and he suggested that this might be one way in which children might start to develop these categories. Subsequent studies found a similar results for French (Chemla, Mintz, Bernal & Christophe, 2009), but the method worked much less well for Dutch (Erkelens, 2009) or German (Stumper, Bannard, Lieven & Tomasello, 2011). This seems to be due to the somewhat more varied word order of Dutch and German and the use of some forms as both determiners and pronouns (i.e. across categories), highlighting the importance of testing proposed learning mechanisms on typologically varied languages. This type of research reveals the structure of information in the input, but it does not tell us whether or how children utilize this information. Although studies have shown that infants can extract structure quite readily from small artificial languages (e.g. Saffran, Aslin & Newport, 1996; Gerken, 2006), putting the two methods together to arrive at a more realistic model is not easy. First, there is the challenge of determining the multiple factors involved in structuring the input. Interestingly, subsequent studies have shown that much stronger results for lexical categories can be obtained if multiple sources of information are used: e.g. distributional information together with phonological information (Monaghan, Chater & Christiansen, 2005). There have been a number of path-breaking computational studies concerned with the learning of syntactic structure. These can take very different approaches to what is built into the model. To give just three examples: in Bod (2009), the input is already syntactically parsed and the model has an inventory of possible syntactic trees. The results show that the learning model moves from item-based to more abstract constructions. A very different approach is taken by Bannard, Lieven, and Tomasello (2009), which uses a Bayesian procedure to extract item-based grammars from children's corpora and shows that these grammars are successful at parsing unique utterances from

https://doi.org/10.1017/S0305000914000282 Published online by Cambridge University Press

two hours of subsequent recordings. Chang, Dell, and Bock's (2006) dual path model uses a connectionist learner which performs next word prediction. It does this on the basis of a meaning system based on the concepts in the message, and a sequencing system based on learning the sequential probabilities between words.

A major issue for all theories of language is how children learn structures for which there is no obvious evidence in the input. The classic example is that of long-distance dependencies: e.g. *Is the man who is tall sad?*, in which the relative clause as a whole (*the man who is tall*) is the subject of the sentence. Chomsky (1980, p. 113) argued that children need to understand constituency and movement to solve this problem, and that this could not be learned from the input, because it contains no direct guide to constituency and because children almost never hear sentences of this type. However, using a 'mini' input of questions containing a single auxiliary and declaratives, a very small number of which contained relative clauses, Lewis and Elman (2001) managed to get a simple recurrent network to correctly predict the next word of complex questions without error. Reali and Christiansen (2004) replicated this using a network trained on real child-directed speech. It seems that both models depended on the calculation of relative bigram statistics to avoid making structure dependence errors. More recent studies of long-distance dependencies in *wh*-questions have shown similar results (Dąbrowska, Rowland & Theakston, 2009). The crucial point is that children do not necessarily have to hear a structure in order to learn how to produce it: their knowledge of constituency develops over time on the basis of the earlier learning of simpler structures and of the dependencies between words. The input does not just 'imprint' on the child: it is interpreted through what is already learned and through a developing network of constructions.

Quantifying productivity and abstraction

The issues of early productivity and abstraction are central to our understanding of the nature of children's linguistic representations and how these might change over time. Following on from Braine (1976), a number of studies suggested that young children's early speech is very lexically specific: it is built out of rote-learned utterances and/or low scope, slot-and-frame patterns (Clark, 1974; Tomasello, 1992; Lieven, Pine & Baldwin, 1997). Studies show that a large proportion of children's early novel utterances can be accounted for on the basis of these lexically based patterns. These can be highly productive but contain relatively low-scope categories in the slots (Dąbrowksa & Lieven, 2005; Lieven, Salomo & Tomasello, 2009; Bannard *et al*., 2009). More recently, experimental evidence also supports the idea that these low-scope patterns are stored

https://doi.org/10.1017/S0305000914000282 Published online by Cambridge University Press

and used by children and adults (Bannard & Matthews, 2008; Arnon & Snider, 2010).

The debate between Valian, Solt, and Stewart (2009) and Pine, Freudenthal, Krajewski, and Gobet (2013) on whether children have an early category of determiner addresses the issue of early abstraction directly. The main controversy has been over how to measure the overlap in determiner use with different nouns. This has led to a recognition of the crucial role played by Zipf's Law in attempts to control for frequency distributions in the measurement of overlap scores and, as such, has clearly made measurements of productivity more sophisticated than they were earlier. The crucial controls are to measure productivity only on lexical items and/or inflections used by both the child and the caretaker(s) and to randomly reduce the adult corpus to the same size as the child's. When this is done, young children are systematically less productive with the determiners *the* and *a/an* by comparison with the adults around them. The same principle has been used to investigate children's productivity with inflectional morphology in Spanish and Polish (Aguado-Orea, 2004; Krajewski, Lieven & Theakston, 2012, respectively). These studies, and others testing for productivity using a somewhat different method (e.g. Stoll *et al*., 2012), indicate that children's productivity of these highly frequent and basic forms reaches that of the adults' child-directed speech by around age 3;0.

Another major focus of research concerns argument-structure constructions and the ongoing debate about how early these are represented abstractly. In a word order language like English, transitives share no lexical material: the speaker and listener have to know that word order codes the subject and object of the sentence. Pinker (1984) suggested that this could be solved by innate linking between the semantic category of agent and that of grammatical subject. This was countered by evidence from children's early transitive utterances (e.g. children's first subjects were often not agents; Bowerman, 1990) and from cross-linguistic evidence (ergative languages in which thematic roles align differently with subjects and objects; Bavin & Stoll, 2013). Tomasello (1992) proposed that children do not start out with innate linking rules but with 'verb island' constructions with no connection between the argument roles in the individual constructions, and that, as they develop, children analogize across these verb islands to build an abstract transitive construction. An earlier proposal by MacWhinney (1975, 1982) involving feature cancellation addresses the linking problem in a somewhat similar way. However, the conclusion that children know nothing abstract about transitives until they are approaching the age of 3;0 has been challenged by experiments showing that children aged 1;9 can find the correct match between two transitive sentences containing different novel verbs and one of two scenes with agent and patient

https://doi.org/10.1017/S0305000914000282 Published online by Cambridge University Press

reversed (Gertner, Fisher & Eisengart, 2006; Fernandes, Marcus, Di Nubila & Vouloumanos, 2006; Noble, Rowland & Pine, 2011). Research is now focused on understanding the ways in which they manage this. While it is clear that children know something more abstract about transitives from much earlier than 3;0, they do not seem to operate with a fully schematic transitive construction until about this age and, in some cases, until much older. For instance, it takes them longer to comprehend non-prototypical transitives, e.g. with an inanimate subject, despite the same abstract argument structure; success is developmentally later if the two actions are identical with roles reversed; there are different rates of development across languages as a function of cues in the input (e.g. argument omission, variable word order; Dittmar, Abbot-Smith, Lieven & Tomasello, 2008; Chan, Lieven & Tomasello, 2009). With development, children show evidence of being able to deal with more abstract instances of argument structure constructions, e.g. with novel as opposed to familiar verbs and with a variety of subject and object types (Childers & Tomasello, 2001; Savage, Lieven, Theakston & Tomasello, 2003; Theakston, Maslen, Lieven & Tomasello, 2012).

The precise nature and abstractness of children's early argument structure constructions is still the subject of intense research, but sorting this out has been greatly improved by the development of a range of methods: e.g. preferential looking studies, 'weird word order' studies, and by the use of novel verbs. Priming, eye-tracking, and EEG studies hold out further promise of being able to monitor the types of cue that children use and the scope of their abstractness as these constructions are processed. Finally, focused diary studies can be extremely useful in capturing relatively rare events in children's language learning. Bowerman's (1988) study of her two children's overgeneralizations was an early and superb example but, more recently, Rowland (2007) on *wh*-questions and Naigles, Hoff, and Vear (2009) on early verbs have shown how useful this method can be.

Explaining errors

It has always been recognized that systematic errors of omission and commission are important because they can be very informative about the child's current linguistic representations. One of the major developments in child language research has been to subject the analysis of errors to rigorous quantitative analysis. In some cases, this has resulted in rethinking the causes of these errors. For instance, although children are frequently reported as making relatively few errors, particularly in inflectional morphology, this is often because the errors are being analyzed at the level of abstract linguistic categories. When analyzed at the level of particular forms, major differences are often found, suggesting that children are learning these forms on a piecemeal basis and gradually abstracting across

https://doi.org/10.1017/S0305000914000282 Published online by Cambridge University Press

the system (Pizzuto & Caselli, 1992; Aguado-Orea, 2004; Pine, Rowland, Lieven & Theakston, 2005; Dąbrowska, 2005; Krajewski, Theakston, Lieven & Tomasello, 2011).

One approach to children's systematic errors is to see them as a reflection of an initially abstract linguistic system in which a parameter is not yet, or wrongly, set. The alternative is that many errors arise from the entrenchment of high-frequency strings which then compete with differential rates of semantic and syntactic generalization in related parts of the network. Studies suggest that children are less likely to make errors with question frames that are frequent in the input (Rowland & Pine, 2000; Rowland, 2007). Non-inversion errors occur on low-frequency strings and children show alternations between correct and incorrect inversion while the system is developing. In an experimental study in which children had to produce questions using combinations of *wh*-words and auxiliaries, Ambridge, Rowland, Theakston, and Tomasello (2006) showed that lexically specific strings of *wh*-words and an auxiliary could account better for the pattern of errors than either the *wh*-word or the auxiliary type alone. On the other hand, entrenched high-frequency strings can also lead to error. Ambridge and Rowland (2009) found that if the children had an entrenched positive question frame such as *What does X?*, they were significantly more likely to make auxiliary doubling errors when the related negative question was elicited (i.e. to say *What does* (+) *she doesn't like*? instead of *What doesn't she like?*).

Another area that has been a focus for both sides of the generativist–usage-based debate is that of optional infinitive (OI) errors – utterances that lack finiteness marking (e.g. *He going, He do that*). In probably the most influential generativist approach to this issue, the Agreement–Tense Omission model (ATOM), Wexler (1998) suggested that while children have correctly set the tense and agreement parameters of their language from a very early stage, they are subject to a unique checking constraint in early development which means one of these features may be optionally underspecified. This theory makes the prediction that, in languages in which subjects can be dropped ('pro-drop' languages), the rate of OI errors will be very low because children do not have to deal with checking the agreement feature on the subject. In support of this, young Spanish-speaking children make few OI errors while German- and Dutch-speaking children, who are learning non-pro-drop languages, show very high rates, with Dutch rates even higher than German. However, Freudenthal, Pine, Aguado-Orea, and Gobet (2007) have shown that differential rates in OI errors can be accounted for by the relative frequency of utterance final, non-finite verbs (Dutch 87%, German 66%, Spanish 26%), which results from the verb-second rule in Dutch and German (see also Wijnen, Kempen & Gillis, 2001). The suggestion is that the child's

https://doi.org/10.1017/S0305000914000282 Published online by Cambridge University Press

processing mechanism may be differentially picking up forms at the ends of utterances, and it is this that gives rise to the different rates of OI errors. In a more recent study, which compared the usage-based position on OI errors with the 'variable learning model' of Legate and Yang (2007), Freudenthal, Pine, and Gobet (2010), found that the particular verbs with which children were more likely to make OI errors were those that occurred more frequently in complex verb phrases in the input, suggesting a specific lexical effect on learning. Theakston, Lieven, and Tomasello (2003) showed a similar effect of input in an experimental study in which children who only ever heard novel verbs in complex verb phrases (i.e. without tense marking) were significantly more likely to produce them as OI errors. However, in support of the idea that children are engaged in gradual abstraction, the results indicated that these children's learning of finiteness marking was developing: if they heard the verb in the finite form, they never incorrectly produced it as a non-finite. Again, development is gradual and piecemeal: children are not leaping from rote-learned utterances to a fully abstract adult grammar!

A second prediction of the ATOM is that accusative-for-nominative errors (e.g. *me do it*) occur when children check tense on the verb and do not therefore also mark case on the subject but use the default form (in English, the accusative, hence *me*-for-*I* errors; see Rispoli, 1994). Kirjavainen, Theakston, and Lieven (2009) found that complex utterances in the input (*Let me do it*) could be the source of these errors. Similarly to the Freudenthal *et al*. (2010) study above, the verbs that children produced in *me*-error utterances appeared in complex sentences containing *me* in the input significantly more often than verbs that the children used correctly with *I*. For at least one of the children, this becomes a productive pattern. She uses the *me* +*V* pattern with finite verbs which are very unlikely to have occurred after *me* in her input (e.g. *me got* . . . , *me can't* . . .). This indicates a process of abstraction from a set of strings learned from the input. These will eventually become of very low, or non-existent, frequency as the frequency of competing, correct, forms builds up.

A problem that has received continuous attention over the last forty years is how children recover from overgeneralization errors (see Ambridge & Lieven, 2011, for a detailed discussion). This is a particularly hard problem because, on the one hand, children must be able to form generalizations in order to use language productively. On the other, they are unlikely to be able to retreat from incorrect generalizations purely on the basis of corrective evidence in the input, particularly since there is often not clear competition between different forms – unlike the case with the *me*-for-*I* errors discussed above. Pinker's (1989) solution to the problem was based on the interaction of verb meaning with syntax. He suggested that children start with broad-range rules that allow a wide range of verbs to appear in a

https://doi.org/10.1017/S0305000914000282 Published online by Cambridge University Press

particular syntactic structure (e.g. action verbs in a causative transitive). Subsequently, children start to form narrow-range semantic verb classes that define which verbs can alternate and which cannot. An alternative, though, as it turns out, not conflicting, line of research is based on the finding that frequency is very important in determining which verbs children are likely to overgeneralize, with consistent results showing that low-frequency verbs are much more vulnerable (Brooks, Tomasello, Dodson & Lewis, 1999; Theakston, 2004). Recent research has shown that both frequency and semantics are independently important and that, rather than verb classes, the issue in semantics is the degree of fit with the construction as a function of what the child has already learned (for a review see Ambridge, Pine, Rowland, Chang & Bidgood, 2013). This is an area which is also being clarified by computational modelling (Alishahi & Stevenson, 2008; Parisien & Stevenson, 2010; Perfors, Tenenbaum & Wonnacott, 2010).

OUTSTANDING ISSUES

Construction-based learning

Theoretical approaches which see the state of the adult language system as an inventory of constructions have been very influential in usage-based approaches to acquisition (Langacker, 1987; Goldberg, 1995). There are two reasons for this. First, constructions are defined as pairings of form and meaning, which contrasts with any approach that treats meaning as separate from the syntactic assembly of sentences. This allows communicative function to be foundational in language learning. Second, constructions can exist at any level of abstraction from the concrete (*He hit me*) through schemas with abstract slots (*X hit me*) to fully abstract constructions such as ($\text{NOUN}_{\text{subject}}$ $\text{VERB}_{\text{causative}}$ $\text{NOUN}_{\text{object}}$). This means that children could build up their inventories of constructions, starting with low-scope, lexically specific schemas and developing more abstract constructions over time. In addition, constructions at different levels of abstractness continue to co-exist in people's inventories. Goldberg (1995) and Verhagen (2005) are two important attempts to characterize the adult inventory and the links between constructions. The general principle of an inventory of constructions at different levels of abstraction fits well with the evidence for both early lexically specific patterns and developing abstraction outlined above. However, much more detailed research is needed on the relationships between constructions and how children's inventories are built. Possibilities are following up the ideas of Tomasello (2003) on analogy, MacWhinney (1982) on feature cancellation, or Goldberg (1995) on inheritance hierarchies. Another is to trace the meanings associated with different related forms and how these become connected in development.

https://doi.org/10.1017/S0305000914000282 Published online by Cambridge University Press

Theakston, Lieven, Pine, and Rowland's (2002) analysis of the different meanings that children initially associate with the different forms of the verb GO (e.g. *going*, *goes*, *went*, *gone*) is an example. Yet another possibility is to explore the ways in which emerging constructions may interact in the course of learning, e.g. Abbot-Smith and Behrens' (2006) analysis of the development of passive and future constructions in German, in which they argued that previously learned constructions could either support or inhibit the order in which the child in their study learned the new constructions. While highly entrenched forms allow the child to show early productivity, they can also negatively impact on later development if the form–meaning mappings that are already learned conflict with those of the new construction (e.g. it is more difficult for children to use familiar verbs in passive constructions if they are already highly entrenched in the active; Dittmar, Abbot-Smith, Lieven & Tomasello, 2014; see also Brandt, Verhagen, Lieven & Tomasello, 2011, on sentential complements). In this context, Chang *et al.*'s (2006) model of sentence production, which showed that early in learning the model interpreted intransitive sentences as transitives due to the entrenchment of the first noun in the thematic role of ACTOR, is a powerful demonstration of how we might be able to model the relationship between constructions in development.

Finally, the analysis of discourse sequences to assess children's linguistic representations could be more exploited. There has been some work on the information that caretakers provide about the relationship between forms and constructions in 'variation sets' or 'replacement sequences' (Küntay & Slobin, 2002), and whether children pick up this information and learn from it (Chouinard & Clark, 2003). But children's rephrasing of either the caretakers' utterances or their own can indicate their understanding of the relationship between forms and constructions as well as, for example, their grasp of missing items in ellipsis.

Building on preverbal development

Infancy research has yielded spectacular findings on preverbal infants' cognitive skills and semantic and pragmatic sensitivities, but much more work is needed to see how these underpin, or are translated into, language development. Research on the development of early phonological categories has been important in this respect (e.g. Pierrehumbert, 2003; Swingley, 2005), and a very promising recent development is research on infants' speed of language processing and how this impacts on their subsequent language development (Fernald & Marchman, 2012). The question of universals and particulars in the development of meaning is also extremely important. We know that infants are sensitive to the semantic categories of the language they are learning from very early on (Choi & Bowerman,

https://doi.org/10.1017/S0305000914000282 Published online by Cambridge University Press

1991; Brown, 1998). On the other hand, this must, presumably, be built on some universal and possibly primate-wide distinctions, for instance between animate and non-animate motion. An interesting attempt to reconcile these findings is Göksun, Hirsh-Pasek, and Golinkoff's (2010) paper which reviews cross-linguistic research on children's understanding of relational terms in spatial events and how this is related to their language learning.

A problem that often makes infancy research difficult to connect to later development is that experiments showing infant sensitivity to a particular discrimination (e.g. in a false belief experiment at 1;1; Onishi & Baillargeon, 2005) do not, of course, mean that the child can utilize this information. But even when infants do show the ability to operate prelinguistically with quite sophisticated knowledge, for instance in the series of findings on early socio-cognitive development (e.g. Carpenter, Nagell & Tomasello, 1998; Gergely, Bekkering & Király, 2002; Lizskowski, Carpenter & Tomasello, 2007), it is still a complex matter to work out how such knowledge is reflected in children's language development. An example is that preverbal infants clearly know what is given and what new for their interactants (Tomasello & Haberl 2003), and this is also reflected quite early in their language (Salomo, Lieven & Tomasello, 2010; Hughes & Allen, 2013), but how these early abilities are translated into more complex syntax is still very under-researched.

More complex constructions

The ways in which children learn to correctly understand and use more complex constructions (e.g. the scope of negation and quantification) has been somewhat neglected by researchers working from construction-based theoretical positions (but see Goldberg, 2006). These have been the subject of extensive research from other perspectives but, while there has been some attention to the ways in which pragmatic considerations, in particular, information structure, might account for some of the phenomena, this has not been studied in detail in terms of development. One important, and often neglected, factor is what children actually hear in the input. Most experiments on children's understanding of these structures operate with very complex and artificial sentences that children are very unlikely indeed to hear – or even read. Many adults (though not all; see Street & Dąbrowksa, 2010) will be able to succeed in these types of task, but how they learn to do so may well be built on earlier and simpler learning and this will depend crucially on how these forms are actually used in terms of structure and meaning (Ariel, in press) as well as on children's developing literacy.

Children's language environments

The characterization of the language learning environment is clearly essential for the attempt to explain any aspect of language development. We have

https://doi.org/10.1017/S0305000914000282 Published online by Cambridge University Press

come a long way in the last forty years, partly as a result of improved methods and partly as a result of public access to corpora, of which the CHILDES database is the outstanding example (MacWhinney, 2000). However, even for English, where we have a considerable number of corpora, the density of recording is often not sufficient to allow for the systematic analyses of productivity and errors outlined above. We also do not know much about what children learn in non-dyadic situations, since many of these corpora are of the child alone with his/her caregiver, and this certainly does not adequately characterize the language environments in which even monolingual children learn, to say nothing of the huge number of children who grow up varyingly multilingual.

The situation is much worse for languages other than English and, in particular, for non-Indo-European languages. Since all typically developing children can learn any language, and since languages differ so greatly on a vast range of factors (see Lieven & Stoll, 2009), we must know whether the explanations for language learning that are being developed for English will work in typologically very different languages (cf. Berman, 1988, for Hebrew). The challenges set by Slobin (1982) are still very much with us.

Individual differences

Although the interaction between domain-general cognitive and social development and differences in language learning has been the subject of considerable research, this has perhaps been more focused on children with atypical language development, for instance those diagnosed with autism or specific language impairment, areas which I have not attempted to cover here. However, if, in fact, the differences between these conditions are more dimensional than categorical (Bishop, 2003), then studies of how these children learn language could have an impact of how we understand more typical language development. In particular, we know that children vary greatly in their rate of development (Pan, Rowe, Singer & Snow, 2005), and that not all adults end up with the same language abilities (Miller & Weinert, 1998/2009). Recent research on the causes and effects of processing speed is one promising avenue for explaining these differences (Hurtado, Marchman & Fernald, 2008), but there is a great deal to do to unravel the complexities of interaction between cognitive processing, social-cognitive development, and environmental influences in the development of language.

CONCLUSIONS

Language development – at least beyond the very early two-word stage – is unique to humans. I am continually struck both by children's communicative creativity and by their clear dependence on what they hear. In the last

https://doi.org/10.1017/S0305000914000282 Published online by Cambridge University Press

forty years we have come a considerable way in identifying the factors involved and how to measure them. Working out how these factors interact is the next, and even more difficult, task.

REFERENCES

Abbot-Smith, K. & Behrens, Heike (2006). How known constructions influence the acquisition of other constructions: the German passive and future constructions. *Cognitive Science* **30**(6), 995–1026.

Aguado-Orea, J. J. (2004). The acquisition of morpho-syntax in Spanish: implications for current theories of development. (Unpublished PhD thesis) University of Nottingham.

Alishahi, A. & Stevenson, S. (2008). A computational model of early argument structure acquisition. *Cognitive Science* **32**(5), 789–834.

Ambridge, B. & Lieven, E. V. M. (2011). *Child language acquisition: contrasting theoretical approaches*. Cambridge: Cambridge University Press.

Ambridge, B., Pine, J. M., Rowland, C. F., Chang, F. & Bidgood, A. (2013). The retreat from overgeneralization in child language acquisition: word learning, morphology and verb argument structure. *Wiley Interdisciplinary Reviews: Cognitive Science* **4**, 47–62.

Ambridge, B. & Rowland, C. F. (2009). Predicting children's errors with negative questions: Testing a schema-combination account. *Cognitive Linguistics* **20**(2), 225–66.

Ambridge, B., Rowland, C. F., Theakston, A. & Tomasello, M. (2006) Comparing different accounts of non-inversion errors in children's non-subject wh-questions: 'What experimental data can tell us?' *Journal of Child Language* **33**(3), 519–57.

Ariel, M. (in press). 'Or' constructions: monosemy versus polysemy. In B. MacWhinney, A. Malchukov & E. A. Moravcsik (eds), *Competing motivations*. Oxford: Oxford University Press.

Arnon, I. & Snider, N. (2010). More than words: frequency effects for multi-word phrases. *Journal of Memory and Language* **62**(1), 67–82.

Bannard, C., Lieven, E. & Tomasello, M. (2009). Modeling children's early grammatical knowledge. *Proceedings of the National Academy of Sciences* **106**(41), 17284–89.

Bannard, C. & Matthews, D. (2008). Stored word sequences in language learning – the effect of familiarity on children's repetition of four-word combinations. *Psychological Science* **19**(3), 241–8.

Bavin, E. & Stoll, S. (eds) (2013). *The acquisition of ergativity*. Amsterdam: Benjamins.

Berman, R. (1988). Word-class distinctions in developing grammars. In Y. Levy, I. M. Schlesinger & M. D. S. Braine (eds), *Categories and processes in language acquisition* 45–72. Hillsdale: Lawrence Erlbaum.

Bishop, D. V. M. (2003). Autism and specific language impairment: categorical distinction or continuum? *Novartis Foundation Symposium* **251**, 213–26.

Bod, R. (2009). From exemplar to grammar: a probabilistic analogy-based model of language learning. *Cognitive Science* **33**, 752–93.

Bowerman, M. (1988). The 'no negative evidence' problem: How do children avoid constructing an overly general grammar? In J. A. Hawkins (ed.), *Explaining language universals*, 73–101. Oxford: Blackwell.

Bowerman, M. (1990). Mapping thematic roles onto syntactic functions: Are children helped by innate linking rules? *Linguistics* **28**, 1251–89.

Braine, M. D. S. (1976). Children's first word combinations. *Monographs of the Society for Research in Child Development* **41**(164), 1–104.

Brandt, S., Verhagen, A., Lieven, E. & Tomasello, M. (2011). German children's productivity with simple transitive and complement-clause constructions: testing the effects of frequency and variability. *Cognitive Linguistics* **22**(2), 325–57.

Brooks, P. J., Tomasello, M., Dodson, K. & Lewis, L. B. (1999). Young children's overgeneralizations with fixed transitivity verbs. *Child Development* **70**(6), 1325–37.

https://doi.org/10.1017/S0305000914000282 Published online by Cambridge University Press

Brown, P. (1998). Children's first verbs in Tzeltal: evidence for an early verb category. *Linguistics* **36**(4), 713–53.
Carpenter, M., Nagell, K. & Tomasello, M. (1998). Social cognition, joint attention, and communicative competence from 9 to 15 months of age. *Monographs of the Society for Research in Child Development*, **63** (4, Serial No. 255).
Chan, A., Lieven, E. & Tomasello, M. (2009). Children's understanding of the agent – patient relations in the transitive construction: cross-linguistic comparisons between Cantonese, German and English. *Cognitive Linguistics* **20**(2), 267–300.
Chang, F., Dell, G. S. & Bock, K. (2006). Becoming syntactic. *Psychological Review* **113**(2), 234–72.
Chemla, E., Mintz, T. H., Bernal, S. & Christophe, A. (2009). Categorizing words using 'frequent frames': what cross-linguistic analyses reveal about distributional acquisition strategies. *Developmental Science* **12**(3), 396–406.
Childers, J. B. & Tomasello, M. (2001). The role of pronouns in young children's acquisition of the English transitive construction. *Developmental Psychology* **37**(6), 739–48.
Choi, S. & Bowerman, M. (1991). Learning to express motion events in English and Korean: the influence of language-specific lexicalization patterns. *Cognition* **41**, 83–121.
Chomsky, N. (1980). The linguistic approach. In M. Piatelli-Palmarini (ed.), *Language and learning: the debate between Jean Piaget and Noam Chomsky*, 109–17. Cambridge, MA: Harvard University Press.
Chouinard, M. M. & Clark, E. V. (2003). Adult reformulations of child errors as negative evidence. *Journal of Child Language* **30**, 637–69.
Clark, R. (1974). Performing without competence. *Journal of Child Language* **1**(1), 1–10.
Dąbrowska, E. (2005). Productivity and beyond: mastering the Polish genitive inflection. *Journal of Child Language* **32**, 191–205.
Dąbrowska, E., and Lieven, E. (2005). Towards a lexically specific grammar of children's question constructions. *Cognitive Linguistics* **16**(3), 437–74.
Dąbrowska, E., Rowland, C. & Theakston, A. (2009). The acquisition of questions with long-distance dependencies. *Cognitive Linguistics* **20**(3), 571–97.
Dittmar, M., Abbot-Smith, K., Lieven, E. & Tomasello, M. (2008). German children's comprehension of word order and case marking in causative sentences. *Child Development* **79**, 1152–67.
Dittmar, M., Abbot-Smith, K., Lieven, E. & Tomasello, M. (2014). Familiar verbs are not always easier than novel verbs: how German pre-school children comprehend active and passive sentences. *Cognitive Science* **38**(1), 128–51.
Erkelens, M. A. (2009). *Learning to categorize verbs and nouns: studies on Dutch*. Utrecht: LOT.
Fernald, A. & Marchman, V. A. (2012). Individual differences in lexical processing at 18 months predict vocabulary growth in typically-developing and late-talking toddlers. *Child Development* **83**, 203–22.
Fernandes, K. J., Marcus, G. F., Di Nubila, J. A. & Vouloumanos, A. (2006). From semantics to syntax and back again: argument structure in the third year of life. *Cognition* **100**(2), B10–20.
Freudenthal, D., Pine, J. M., Aguado-Orea, J. & Gobet, F. (2007). Modeling the developmental patterning of finiteness marking in English, Dutch, German, and Spanish using MOSAIC. *Cognitive Science* **31**(2), 311–41.
Freudenthal, D., Pine, J. & Gobet, F. (2010). Explaining quantitative variation in the rate of Optional Infinitive errors across languages: a comparison of MOSAIC and the Variational Learning Model. *Journal of Child Language* **37**(3), 643–99.
Gergely, G., Bekkering, H. & Király, I. (2002). Rational imitation in preverbal infants. *Nature* **415**, 755.
Gerken, L. A. (2006). Decisions, decisions: infant language learning when multiple generalizations are possible. *Cognition* **98**, B67–74.
Gertner, Y., Fisher, C. & Eisengart, J. (2006). Learning words and rules: abstract knowledge of word order in early sentence comprehension. *Psychological Science* **17**(8), 684–91.

https://doi.org/10.1017/S0305000914000282 Published online by Cambridge University Press

Göksun, T., Hirsh-Pasek, K. & Golinkoff, R. M. (2010). Trading spaces: carving up events for learning language. *Perspectives on Psychological Science* **5**, 33–42.
Goldberg, A. E. (1995). *Constructions: a Construction Grammar approach to argument structure*. Chicago: University of Chicago Press.
Goldberg, A. E. (2006). *Constructions at work: the nature of generalization in language*. Oxford: Oxford University Press.
Hughes, M. & Allen, S. E. M. (2013). The effect of discourse-pragmatic features on subject omission in English. *Journal of Pragmatics* **56**, 15–31.
Hurtado, N., Marchman, V. A. & Fernald, A. (2008). Does input influence uptake? Links between maternal talk, processing speed and vocabulary size in Spanish-learning children. *Developmental Science* **11**, F31–9.
Kirjavainen, M., Theakston, A. & Lieven, E. (2009). Can input explain children's me-for-I errors? *Journal of Child Language* **36**(5), 1091–114.
Krajewski, G., Lieven, E. & Theakston, A. (2012). Productivity of a Polish child's inflectional noun morphology: a naturalistic study. *Morphology* **22**(1), 9–34.
Krajewski, G., Theakston, A., Lieven, E. & Tomasello, M. (2011). How Polish children switch from one case to another when using novel nouns: challenges for models of inflectional morphology. *Language and Cognitive Processes* **26**, 830–61.
Küntay, A. C. & Slobin, D. I. (2002). Putting interaction back into child language: examples from Turkish. *Psychology of Language and Communication* **6**, 5–14.
Langacker, R. W. (1987). *Foundations of Cognitive Grammar*, vol. **1**. Stanford: Stanford University Press.
Legate, J. A. & Yang, C. (2007). Morphosyntactic learning and the development of tense. *Language Acquisition* **14**, 315–44.
Lewis, J. D. & Elman, J. L. (2001). Learnability and the statistical structure of language: poverty of stimulus arguments revisited. *Proceedings of the Twenty-Sixth Annual Boston University Conference on Language Development*, 359–70. Somerville, MA: Cascadilla Press.
Lieven, E. V. M., Pine, J. M. & Baldwin, G. (1997). Lexically-based learning and early grammatical development. *Journal of Child Language* **24**(1), 187–219.
Lieven, E. V. M., Salomo, D. & Tomasello, M. (2009). Two-year-old children's production of multiword utterances: a usage-based analysis. *Cognitive Linguistics* **20**(3), 481–507.
Lieven, E. & Stoll, S. (2009). Language development. In M. Bornstein (ed.), *The handbook of cross-cultural developmental science*, 143–60. New York: Psychology Press.
Lizskowski, U., Carpenter, M. & Tomasello, M. (2007). Reference and attitude in infant pointing. *Journal of Child Language* **34**,1–20.
MacWhinney, B. (1975). Rules, rote, and analogy in morphological formations by Hungarian children. *Journal of Child Language* **2**, 65–77.
MacWhinney, B. (1982). Basic syntactic processes. In S. Kuczaj (ed.), *Language development. Vol. 1: syntax and semantics*, 73–136. Hillsdale, NJ: Lawerence Erlbaum.
MacWhinney, B. (2000). *The CHILDES Project: tools for analyzing talk. Vol. 1: transcription format and programs. Vol. 2: the database*. Mahwah, NJ: Lawrence Erlbaum Associates.
Miller, J. & Weinert, R. (1998/2009). *Spontaneous spoken language: syntax and discourse*. Oxford: Oxford University Press.
Mintz, T. H. (2003). Frequent frames as a cue for grammatical categories in child directed speech. *Cognition* **90**, 1–117.
Monaghan, P., Chater, N. & Christiansen, M. H. (2005). The differential contribution of phonological and distributional cues in grammatical categorization. *Cognition* **96**, 143–82.
Naigles, L., Hoff, E. & Vear, D. (2009). Flexibility in early verb use: evidence from a multiple-n diary study. *Monographs of the Society for Research in Child Development* **74**(2), 1–143.
Noble, C. H, Rowland, C. F. & Pine, J. M. (2011). Comprehension of argument structure and semantic roles: evidence from infants and the forced-choice pointing paradigm. *Cognitive Science* **35**(5), 963–82.
Onishi, K. H. & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science* **308**, 255–8.

https://doi.org/10.1017/S0305000914000282 Published online by Cambridge University Press

Pan, B., Rowe, M., Singer, J. & Snow, C. (2005). Maternal correlates of growth in toddler vocabulary production in low-income families. *Child Development* **76**(4), 763–82.
Parisien, C. & Stevenson, S. (2010). Learning verb alternations in a usage-based Bayesian model. In *Proceedings of the 32nd Annual Conference of the Cognitive Science Society*, Portland, Oregon, August, online: <http://mindmodeling.org/cogsci2010/papers/0641/paper0641.pdf>.
Perfors, A., Tenenbaum, J. B. & Wonnacott, E. (2010). Variability, negative evidence, and the acquisition of verb argument constructions. *Journal of Child Language* **37**(3), 607–42.
Pierrehumbert, J. B. (2003). Phonetic diversity, statistical learning and acquisition of phonology. *Language and Speech* **46**, 115–54.
Pine, J. M., Freudenthal, D., Krajewski, G. & Gobet, F. (2013). Do young children have adult-like syntactic categories? Zipf's law and the case of the determiner. *Cognition* **127**(3), 345–60.
Pine, J., Rowland, C. F., Lieven, E. V. M. & Theakston, A. L. (2005). Testing the Agreement/Tense Omission Model: why the data on children's use of non-nominative subjects count against the ATOM. *Journal of Child Language* **32**, 269–89.
Pinker, S. (1984). *Language learnability and language development*. Cambridge, MA: Harvard University Press.
Pinker, S. (1989). *Learnability and cognition: the acquisition of argument structure*. Cambridge, MA: MIT Press.
Pizutto, E. & Caselli, C. (1992). The acquisition of Italian morphology: implications for models of language development. *Journal of Child Language* **19**, 491–557.
Reali, F. & Christiansen, M. H. (2004). Uncovering the richness of the stimulus: structure dependence and indirect statistical evidence. *Cognitive Science* **29**, 1007–28.
Redington, M., Chater, N. & Finch, S. (1998). Distributional information: a powerful cue for acquiring syntactic categories. *Cognitive Science* **22**(4), 425–69.
Rispoli, M. (1994). Pronoun case overextension and paradigm building. *Journal of Child Language* **21**, 157–72.
Rowland, C. F. (2007). Explaining errors in children's questions. *Cognition* **104**(1), 106–34.
Rowland, C. F. & Pine, J. M. (2000). Subject–auxiliary inversion errors and wh-question acquisition: 'what children do know?' *Journal of Child Language* **27**(1), 157–81.
Saffran, J. R., Aslin, R. N. & Newport, E. L. (1996). Statistical learning by 8-month olds. *Science* **274**, 1926–8.
Salomo, D., Lieven, E. & Tomasello, M. (2010). Young children's sensitivity to new and given information when answering predicate-focus questions. *Applied Psycholinguistics* **31**(1), 101–15.
Savage, C., Lieven, E., Theakston, A. & Tomasello, M. (2003). Testing the abstractness of young children's linguistic representations: lexical and structural priming of syntactic constructions? *Developmental Science* **6**(5), 557–67.
Slobin, D. I. (1982). Universal and particular in the acquisition of language. In E. Wanner & L. Gleitman (eds), *Language acquisition: the state of the art*, 128–70. Cambridge: Cambridge University Press.
Stoll, S., Bickel, B., Lieven, E., Banjade, G., Bhatta, T. N., Gaenszle, M., Paudyal, N. P., Pettigrew, J., Rai, I. P., Rai, M. & Rai, N. K. (2012). Nouns and verbs in Chintang: children's usage and surrounding adult speech. *Journal of Child Language* **39**, 284–321.
Street, J. & Dąbrowska, E. (2010). More individual differences in language attainment: How much do adult native speakers of English know about passives and quantifiers? *Lingua* **120**, 2080–94.
Stumper, B., Bannard, C., Lieven, E. & Tomasello, M. (2011). 'Frequent frames' in German child-directed speech: a limited cue to grammatical categories. *Cognitive Science* **35**, 1190–205.
Swingley, D. (2005). 11-months-olds' knowledge of how familiar words sound. *Developmental Science* **8**, 432–43.
Theakston, A. L. (2004). The role of entrenchment in children's and adults' performance on grammaticality judgement tasks. *Cognitive Development* **19**(1), 15–34.

https://doi.org/10.1017/S0305000914000282 Published online by Cambridge University Press

Theakston, A. L., Lieven, E. V. M., Pine, J. M. & Rowland, C. F. (2002). Going, going, gone: the acquisition of the verb 'go'. *Journal of Child Language* **29**, 783–811.
Theakston, A., Lieven, E. & Tomasello, M. (2003). The role of input in the acquisition of third-person singular verbs in English. *Journal of Speech, Language & Hearing Research* **46**, 863–77.
Theakston, A. L., Maslen, R., Lieven, E. V. M. & Tomasello, M. (2012). The acquisition of the active transitive construction in English: a detailed case study. *Cognitive Linguistics* **23**(1), 91–128.
Tomasello, M. (1992). *First verbs: a case study of early grammatical development*. Cambridge: Cambridge University Press.
Tomasello, M. (2003). *Constructing a language: a usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
Tomasello, M. & Haberl, K. (2003). Understanding attention: 12- and 18-month-olds know what's new for other persons. *Developmental Psychology* **39**, 906–12.
Valian, V., Solt, S. & Stewart, J. (2009). Abstract categories or limited-scope formulae? The case of children's determiners. *Journal of Child Language* **36**(4), 743–78.
Verhagen, Arie (2005). *Constructions of intersubjectivity: discourse, syntax, and cognition*. Oxford: Oxford University Press.
Wexler, K. (1998). Very early parameter setting and the unique checking constraint: a new explanation of the optional infinitive stage. *Lingua* **106**(1/4), 23–79.
Wijnen, F., Kempen, M. & Gillis, S. (2001). Bare infinitives in Dutch early child language: An effect of input? *Journal of Child Language* **28**, 629–60.

https://doi.org/10.1017/S0305000914000282 Published online by Cambridge University Press